

## ANALISIS SENTIMEN BERBASIS ASPEK PADA APLIKASI ELEKTRONIK SURVEI KEPUASAN MASYARAKAT (E-SKM) JAWA TENGAH MENGGUNAKAN INDOBERT

Refo Labib Mustofa<sup>\*1</sup>, Tarno<sup>2</sup>, Catur Edi Widodo<sup>3</sup>

<sup>1,2,3</sup>Universitas Diponegoro, Semarang

Email: <sup>1</sup>refolabib@students.undip.ac.id, <sup>2</sup> tarno@lecturer.undip.ac.id, <sup>3</sup> caturediwido@lecturer.undip.ac.id

\*Penulis Korespondensi

(Naskah masuk: 17 April 2025, diterima untuk diterbitkan: 08 Juli 2025)

### Abstrak

Perkembangan tentang *Natural Language Processing* (NLP) semakin berkembang dengan pesat, salah satunya yaitu dalam bidang analisis sentimen. Dalam dunia bisnis, analisis sentimen sangat diperlukan untuk mengetahui dan memahami persepsi pelanggan terhadap produk yang telah didapatkan dari perusahaan. Hal yang sama juga berlaku pada sektor pemerintahan. Pemerintah sebagai penyelenggara pelayanan publik harus dapat mengetahui persepsi dari pengguna layanan terhadap penyelenggaraan pelayanan publik tersebut sebagai bahan perbaikan kualitas layanan. Aplikasi E-SKM merupakan aplikasi milik Pemerintah Provinsi Jawa Tengah yang saat ini hanya mengolah nilai survei layanan meliputi sembilan aspek pertanyaan, sedangkan data saran/masukan pada aplikasi ini belum dimanfaatkan lebih lanjut. Pada penelitian ini, dilakukan analisis sentimen pada data saran/masukan tersebut untuk menggali informasi tambahan yang dapat meningkatkan pemahaman pemerintah terhadap kepuasan pengguna layanan. Metode yang diusulkan yaitu menggunakan pendekatan analisis sentimen berbasis aspek menggunakan model IndoBERT. Pendekatan berbasis aspek ditujukan agar dapat diketahui aspek apa saja yang paling banyak dibicarakan oleh pengguna layanan, terutama yang berhubungan dengan sembilan aspek pertanyaan tersebut. Pada penelitian ini juga digunakan kamus leksikon sebagai pelabelan data, kemudian pendekatan berbasis aturan (*rule-based*) digunakan dalam proses klasifikasi aspek yang berkaitan dengan sembilan aspek pertanyaan. Selain itu, penelitian ini bertujuan untuk mengukur kemampuan model IndoBERT dalam proses klasifikasi sentimen dengan beberapa skenario yang berbeda. Dari hasil analisis, model evaluasi IndoBERT berjalan dengan baik. Hal ini dilihat dari nilai rata-rata parameter evaluasi seperti akurasi, *precision*, *recall*, dan *F1-score* mencapai 95%. Penerapan model ini memiliki kontribusi pada data aplikasi E-SKM untuk mendapatkan informasi sentimen dan aspek pada data pelayanan publik di pemerintahan yang dapat digunakan sebagai bahan pengambilan keputusan pada level manajemen kebijakan.

**Kata kunci:** NLP, analisis sentimen, ABSA, IndoBERT

## ASPECT-BASED SENTIMENT ANALYSIS ON THE CENTRAL JAVA PUBLIC SATISFACTION SURVEY ELECTRONIC APPLICATION (E-SKM) USING INDOBERT

### Abstract

The field of *Natural Language Processing* (NLP) is rapidly advancing, particularly in sentiment analysis. In the business world, sentiment analysis is essential for understanding customer perceptions of products they have received from a company. The same applies to the government sector, where it is crucial for public service providers to gain insight into user perceptions of public services as a basis for service improvement. The E-SKM application, owned by the Central Java provincial government, currently processes only service survey scores covering nine question aspects, while suggestions/feedback data from this application have not yet been fully utilized. In this study, sentiment analysis was conducted on the suggestion/feedback data to extract additional insights that could improve understanding of user satisfaction. The proposed method involves an aspect-based sentiment analysis approach using the IndoBERT model. This aspect-based approach aims to identify the aspects most frequently mentioned by service users, particularly those related to the nine survey aspects. A lexicon-based approach was used for data labeling, followed by a rule-based approach for classifying aspects associated with the nine questions. Additionally, this study aims to assess the performance of the IndoBERT model in sentiment classification across several scenarios. Evaluation results indicate that IndoBERT performs well, with average metrics such as accuracy, precision, recall, and F1-score reaching 95%. The implementation of this model contributes to the E-SKM application data by providing sentiment and aspect information on public service data within the government, which can be used as a basis for decision-making at the policy management level.

**Keywords:** NLP, sentiment analysis, ABSA, IndoBERT

## 1. PENDAHULUAN

Kepuasan pelanggan menjadi salah satu indikator kunci dalam mengevaluasi keberhasilan produk atau layanan (Elifneh, et al., 2020). Survei kepuasan pelanggan secara signifikan dapat meningkatkan loyalitas pelanggan, di mana perusahaan yang mengelola hubungan dengan pelanggan secara efektif menunjukkan retensi yang lebih tinggi dan peningkatan kinerja keuangan (Diekson et al., 2023). Selain itu, perusahaan yang menggunakan pendekatan berbasis data untuk mengukur kepuasan pelanggan cenderung membuat keputusan yang lebih baik, terutama dalam menghadapi dinamika pasar yang berubah dengan cepat (Guerola et al., 2021). Survei kepuasan pelanggan tidak hanya memberikan gambaran jelas tentang apa yang perlu diperbaiki, tetapi juga membantu mengidentifikasi peluang untuk inovasi produk yang mungkin tidak terlihat oleh manajemen internal. Dalam pemerintahan, survei kepuasan masyarakat memainkan peran penting dalam meningkatkan kualitas layanan publik (Romero et al., 2022). Survei kepuasan pelanggan juga berperan penting dalam meningkatkan kepercayaan publik terhadap layanan pemerintah, yang pada gilirannya memperkuat legitimasi dan kredibilitas institusi publik (Citra et al., 2021). Survei kepuasan juga membantu pemerintah memahami ekspektasi masyarakat terhadap pelayanan publik, sehingga kebijakan yang dihasilkan lebih tepat sasaran dan meningkatkan akuntabilitas serta transparansi layanan.

Saat ini Pemerintah Provinsi Jawa Tengah memiliki aplikasi Elektronik Survei Kepuasan Masyarakat (E-SKM) untuk melaksanakan Survey Kepuasan Masyarakat (SKM) yang dapat digunakan semua perangkat daerah untuk melakukan survei kepuasan masyarakat. Sejak digunakan mulai pada Tahun 2019, penggunaan aplikasi E-SKM menghasilkan data pada tabel data survei terdapat 259921 baris data. Data saran/masukan yang memiliki jumlah cukup besar pada aplikasi E-SKM menjadi permasalahan utama, karena pada aplikasi E-SKM saat ini belum dapat mengolah data tersebut menjadi informasi yang bermanfaat. Jika hanya mengandalkan metode tradisional seperti survei atau kuesioner, sering kali menunjukkan keterbatasan yang signifikan (Ibnu et al., 2024). Pendekatan ini sering dianggap tidak efisien, kurang transparan, dan cenderung menghasilkan respons umpan balik yang terbatas.

Dari data saran/masukan yang memiliki jumlah yang banyak ini memerlukan sebuah pendekatan analisis yang kompleks untuk mendapatkan sebuah wawasan atau informasi yang terkandung pada setiap data penggunaannya (Hariri et al., 2019). Dalam konteks ini, analisis sentimen muncul sebagai pendekatan yang dapat memberikan pemahaman mendalam terhadap persepsi dan emosi pengguna (Zhang et al., 2020). Analisis sentimen telah

menyebarkan ke hampir setiap industri, termasuk layanan, keuangan, *e-commerce*, barang konsumen, telekomunikasi, perawatan kesehatan, kampanye politik, acara sosial, dan pemilihan (Wankhade et al., 2024). Namun analisis sentimen sering mengabaikan hubungan sentimen antara istilah aspek dan sentimen yang diekspresikan oleh kalimat ulasan itu sendiri (Gu et al., 2025). Dalam hal ini ada sebuah pendekatan analisis sentimen dengan berbasis aspek guna melakukan ekstraksi sentimen yang lebih mendalam (Brauwiers & Frasinca, 2021).

Analisis sentimen berbasis aspek tidak hanya berguna untuk menangkap makna dan emosi dari saran/masukan pengguna, tetapi juga dapat mengekstraksi aspek-aspek spesifik yang relevan dari data tersebut. Analisis sentimen berbasis aspek memberikan keunggulan dalam memahami opini publik secara lebih rinci dan *actionable*. Alih-alih hanya mengidentifikasi sentimen positif, negatif, atau netral secara umum, pendekatan ini memungkinkan penggalian sentimen terhadap atribut atau fitur spesifik dari suatu produk, layanan, atau kebijakan (Jadhav & Sonar, 2021). Metode analisis sentimen berbasis aspek dengan leksikon juga dapat digunakan sebagai contoh dalam menganalisis data ulasan aplikasi pintar (*smart app*) milik pemerintah United Emirate Arab (UEA) (Alqaryouti et al., 2024).

Dalam konteks E-SKM, analisis sentimen berbasis aspek dapat mengungkap area spesifik layanan publik yang membutuhkan perhatian dan perbaikan. Misalnya, publik mungkin merasa puas dengan kemudahan akses layanan online, namun tidak puas dengan responsivitas petugas terkait keluhan. Informasi terinci ini, yang sulit didapatkan melalui pendekatan sentimen analisis konvensional, memungkinkan pemerintah untuk mengoptimalkan alokasi sumber daya dan merancang intervensi yang lebih terarah dan efektif (Ma et al., 2022). Lebih lanjut, dengan mengidentifikasi aspek-aspek spesifik yang memengaruhi kepuasan masyarakat, pemerintah dapat memprioritaskan perbaikan dan mengembangkan strategi yang lebih responsif terhadap kebutuhan dan harapan publik (Ziemski et al., 2020). Analisis sentimen berbasis aspek juga berperan penting dalam mendeteksi isu-isu *emerging* dan tren sentimen terkait aspek-aspek tertentu dari layanan publik, sehingga memungkinkan antisipasi dan respon proaktif dari pemerintah (Chaudhary et al., 2022).

Dari beberapa penelitian sebelumnya terkait analisis sentimen berbasis aspek juga digunakan dalam tugas klasifikasi ataupun prediksi sentimen dengan menggunakan beberapa model algoritma seperti *Support Vector Machine* (SVM), Naive Bayes dan algoritma yang lain. Salah satu model terbaru yang dapat digunakan dalam analisis sentimen adalah menggunakan model BERT (*Bidirectional Encoder Representations from Transformers*), sebuah model transformer yang memiliki kemampuan pemahaman konteks dan hubungan antar kata yang sangat baik

(Devlin, 2018). Penerapan algoritma BERT dalam analisis sentimen pada aplikasi survei didasarkan pada temuan penelitian-penelitian terkini yang menunjukkan keunggulan model ini dalam memahami konteks dan kompleksitas bahasa. BERT juga berhasil mengatasi keterbatasan model sekuensial tradisional dengan menggunakan metode *pre-training* pada data teks besar, memungkinkan pemahaman kata dan kalimat dalam konteks yang lebih luas dan kontekstual (Pinto, 2023).

Model BERT ini menggunakan data pelatihan berupa korpus dan Wikipedia dalam Bahasa Inggris. Sedangkan dalam penelitian ini data yang digunakan adalah menggunakan Bahasa Indonesia, sehingga pada proposal penelitian ini akan menggunakan IndoBERT. IndoBERT adalah model bahasa pra-latih berdasarkan arsitektur BERT yang dirancang khusus untuk bahasa Indonesia (Koto, 2020). IndoBERT dilatih menggunakan teks dari berbagai sumber dalam bahasa Indonesia, termasuk Wikipedia dan teks dari situs web berita. Pada beberapa penelitian sebelumnya BERT unggul dalam tugas analisis sentimen jika dibandingkan dengan model algoritma *machine learning* dan *deep learning* (Selvakumar & Lakshmanan, 2022). Pada penelitian lain dalam tugas analisis komentar negatif dengan mengelompokkan ke dalam tipe komentar, model IndoBERT juga menunjukkan performa yang cukup baik jika dibandingkan dengan model BERT lainnya seperti Multilingual-BERT dan IndoRoBERTa (Nabilah et al., 2023). Selain itu, model BERT seperti DeBERTa juga dapat dikombinasikan dengan *Gated Recurrent Unit* (GRU) untuk melakukan analisis sentimen (Assiri et al., 2024). Hasil dari DeBERTa dan GRU dengan konfigurasi tersebut menghasilkan akurasi tertinggi yaitu 96.87%. Namun dari hasil ini belum diketahui jika menggunakan variasi pada konfigurasi saat pelatihan data.

Pada proposal penelitian ini akan difokuskan pada analisis sentimen berbasis aspek terhadap data saran/masukan dari aplikasi E-SKM Pemerintah Provinsi Jawa Tengah menggunakan IndoBERT dengan tujuan dapat memperoleh informasi sentimen dan aspek yang dapat digunakan sebagai bahan pengambilan keputusan dalam kebijakan pelayanan publik di pemerintahan. Selain itu, evaluasi model IndoBERT pada tugas ini juga dapat dijadikan referensi untuk penelitian selanjutnya pada model IndoBERT. Pada penelitian ini juga menghasilkan *dataset* ulasan dari survey pelayanan publik beserta sentimennya yang dapat berkontribusi pada penelitian selanjutnya pada bidang NLP.

## 2. METODE PENELITIAN

### 2.1. Dataset

Bahan Penelitian ini menggunakan data saran/masukan responden dari aplikasi ESKM yang dikumpulkan dalam rentang waktu Tahun 2019 hingga Semester 1 Tahun 2024. Data ini tersimpan

dalam *database* MySQL dengan total ukuran sekitar 1,1 Gigabyte (GB). Fokus penelitian ini adalah pada tabel 'utama\_survei', yang berisi 277,338 baris data. Data yang akan dianalisis adalah teks pada kolom 'masukan', yang merupakan jenis data string.

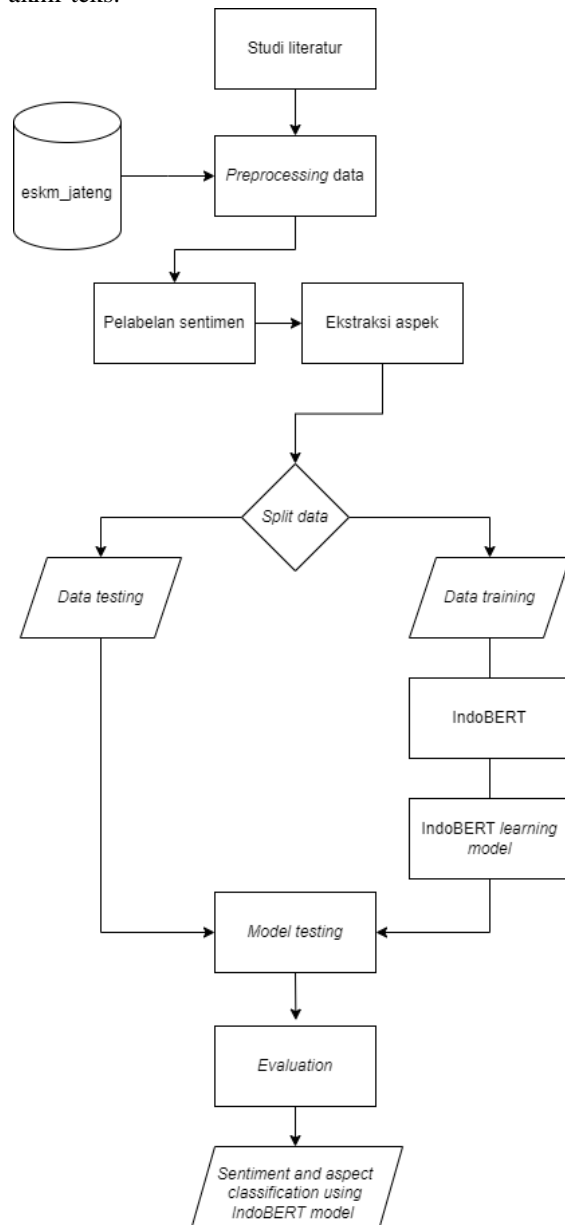
Untuk keperluan analisis sentimen, data tersebut akan dibagi menjadi tiga variasi perbandingan menggunakan model IndoBERT, yaitu 80:20, 70:30, dan 60:40. Pembagian ini bertujuan untuk menguji dan membandingkan performa model dalam berbagai skenario, sehingga dapat menentukan perbandingan yang paling efektif untuk tujuan analisis. Salah satu bahan utama yang digunakan dalam penelitian ini adalah IndoBERT.

### 2.2. Alur Kerja Penelitian

Prosedur penelitian ini terdiri dari beberapa tahapan yang terstruktur dan sistematis. Tahapan-tahapan tersebut meliputi studi literatur yang komprehensif untuk membangun landasan teori yang kuat. Pengambilan data pada *database* pada aplikasi E-SKM hanya fokus pada data saran/masukan dari hasil survei pada pertanyaan terbuka. Selanjutnya, data yang telah dikumpulkan akan dilakukan pelabelan data sentimen menggunakan kamus leksikon Indonesia dan ekstraksi aspek menggunakan metode *rule-based*. Pada proses pelabelan sentimen ini dilakukan secara otomatis menggunakan Python dengan mencocokkan pada kamus leksikon yang sudah memiliki label sentimen pada tiap kata. Pada penelitian ini tidak lagi menggunakan validator dalam mengolah data sentimen dikarenakan jumlah data yang besar, selain itu penggunaan kamus leksikon juga merujuk pada hasil penelitian sebelumnya yang sudah dilakukan validator manusia dalam bidang bahasa. Data yang telah diolah kemudian akan dibagi menjadi data latih dan data uji untuk keperluan pelatihan model analisis sentimen. Pelatihan model ini bertujuan untuk melakukan klasifikasi dan penentuan aspek sentimen yang terdapat dalam data. Tahap selanjutnya adalah pengujian dan evaluasi model yang telah dilatih untuk mengukur performanya dalam menganalisis sentimen. Pada tahapan pengolahan data, penelitian ini akan menggunakan bahasa pemrograman Python beserta *library-library* yang umum digunakan dalam pemrosesan bahasa alami, seperti NLTK, spaCy, dan TensorFlow. *Library-library* ini dipilih karena kemampuannya dalam menangani tugas-tugas NLP, termasuk analisis sentimen. Secara keseluruhan, alur kerja penelitian ini, yang divisualisasikan pada Gambar 1.

Data yang akan digunakan akan dilakukan proses filter dan *preprocessing*. Pada proses filter akan diterapkan dengan menghilangkan data yang memiliki nilai null atau kosong pada kolom 'masukan'. Kemudian pada tahap *preprocessing* dilakukan proses *cleansing* meliputi mengubah teks menjadi huruf kecil, menghilangkan angka, menghilangkan tanda baca, mengganti spasi berlebih

dengan satu spasi, menghilangkan spasi di awal dan akhir teks.



Gambar 1. Alur kerja penelitian

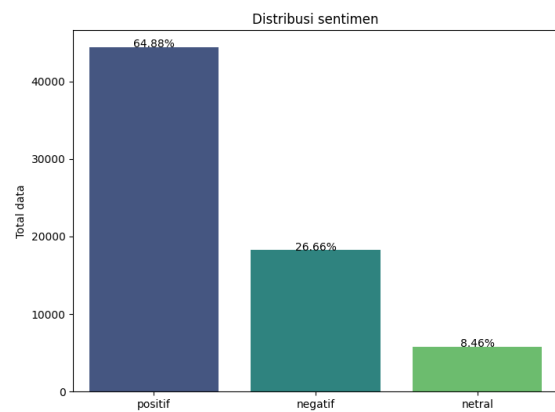
Penggunaan *stop word* (penghilangan kata) tidak dilakukan dalam proses *preprocessing* karena seringkali justru menghilangkan konteks dalam memahami kalimat utuh (Hidayat et al., 2023). Kemudian pada proses pelabelan data sentimen dilakukan dengan menggunakan kamus leksikon bahasa Indonesia. Selain melakukan pelabelan sentimen, dilakukan ekstraksi aspek yang terkait dengan sembilan aspek pertanyaan pada kuisioner survei yaitu persyaratan, prosedur, kecepatan, biaya, pelayanan, kompetensi petugas, perilaku petugas, penanganan pengaduan, sarana prasarana.

Pada penelitian ini, proses pelatihan model IndoBERT dilakukan dengan beberapa skenario. Pada skenario pertama, dilakukan pelatihan data menggunakan model IndoBERT dengan beberapa

variasi perbandingan yaitu 80:20, 70:30 dan 60:40. Skenario ini bertujuan mengukur kemampuan model dari penggunaan data latih yang besar sampai dengan data latih yang kecil, apakah terjadi overfitting dalam pelatihan sehingga mempengaruhi performa model IndoBERT. Dari hasil skenario pertama, rasio perbandingan dalam proses pelatihan data yang menghasilkan performa terbaik akan diterapkan pada skenario kedua, yaitu melakukan pelatihan model IndoBERT dengan berbagai variasi jumlah data. Dalam penelitian ini akan menggunakan 3 sampel data dengan jumlah yang berbeda, yaitu data pada sektor sosial yang berjumlah 12146 data, data pada sektor transportasi sebanyak 10769 data dan data pada sektor kesehatan sejumlah 8871 data.

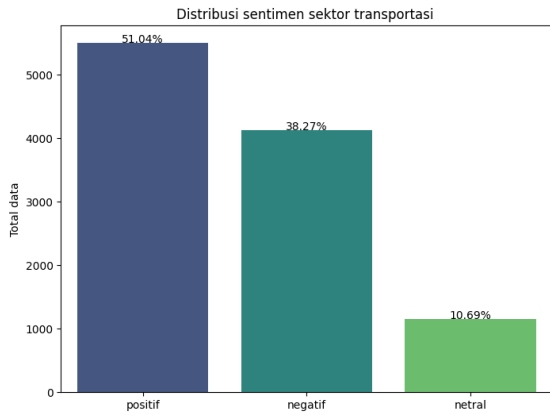
### 3. HASIL DAN PEMBAHASAN

Data yang sudah dilakukan proses pelabelan menggunakan kamus leksikon menghasilkan sentimen positif sebanyak 64.88%, sentimen negatif 26.66% dan sentimen netral 8.46%. Dari hasil ini kemudian dilakukan *breakdown* pada beberapa sampel data seperti pada sektor sosial, sektor transportasi dan sektor kesehatan. Hal ini bertujuan untuk mendapatkan gambaran lebih rinci kontribusi data pada tiap kelas sentimen. Jika dilihat dari keseluruhan data seperti pada Gambar 2, sentimen positif memiliki persentase paling banyak dan memiliki rentang yang jauh dengan sentimen negatif dan netral.

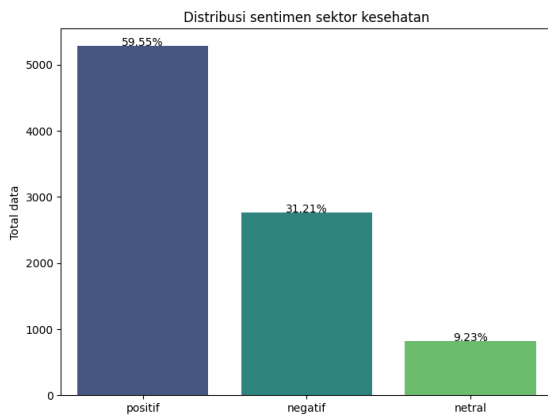


Gambar 2. Distribusi sentimen

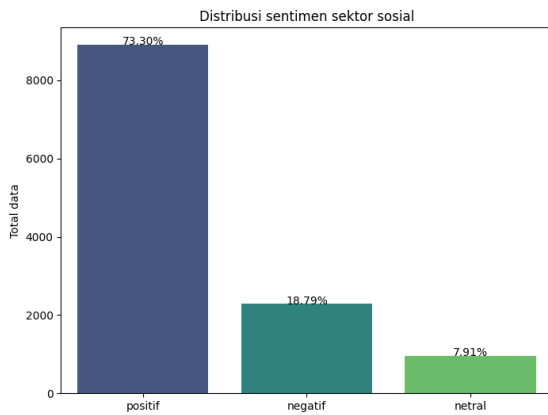
Tetapi ketika dilakukan *breakdown* dengan sampel data pada beberapa sektor seperti sektor perhubungan, sektor sosial dan sektor kesehatan, maka akan terlihat perbedaan yang cukup signifikan pada kelas sentimen. Sektor yang paling terlihat adalah sektor perhubungan yang bisa dilihat pada Gambar 3, terlihat kenaikan sentimen negatif yang cukup banyak jika dibandingkan dengan distribusi sentimen secara keseluruhan. Hal yang sama juga dapat dilihat pada sektor kesehatan seperti pada Gambar 4. Sedangkan pada aspek sosial seperti pada Gambar 5, distribusi sentimen positif lebih banyak dari kelas sentimen negatif dan netral.



Gambar 3. Distribusi sentimen sektor transportasi



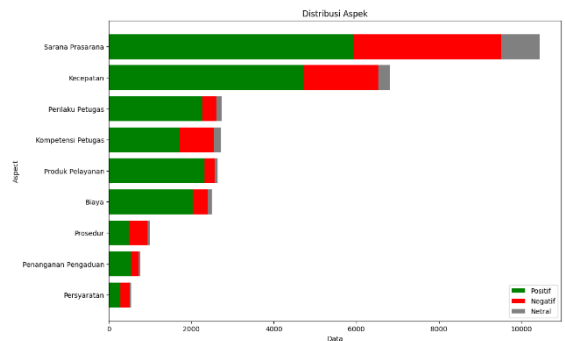
Gambar 4. Distribusi sentimen sektor kesehatan



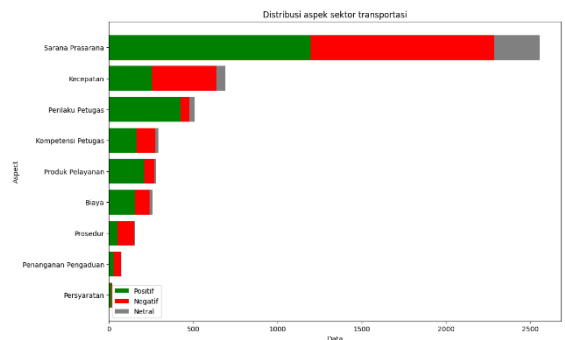
Gambar 5. Distribusi sentimen sektor sosial

Pada tahapan ekstraksi aspek dilakukan dengan pendekatan berbasis aturan (*rule-based*) dengan mengekstraksi kata-kata yang berkaitan dengan sembilan aspek pertanyaan pada kuesioner survei. Kemudian hasil dari ekstraksi aspek yang berkaitan dengan sembilan aspek pada kuisisioner survei, didapatkan aspek terkait sarana prasarana menjadi aspek yang paling banyak dibicarakan dibandingkan dengan aspek yang lain seperti pada Gambar 6. Kemudian aspek persyaratan menjadi aspek yang paling sedikit menjadi masukan pada kuisisioner survei. Pada ekstraksi aspek ini juga dilakukan hal sama seperti pada pelabelan sentimen dengan melakukan *breakdown* pada masing-masing sampel sektor. Jika dilihat pada hasil *breakdown* seperti pada

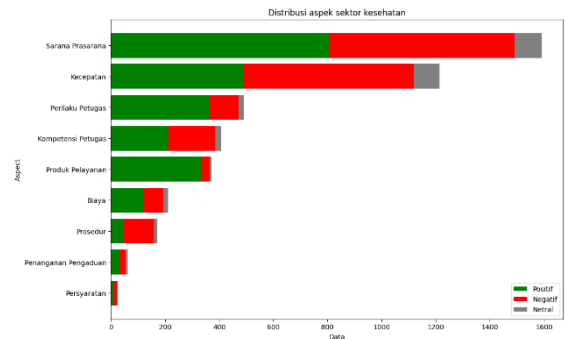
Gambar 7, Gambar 8 dan Gambar 9, aspek sarana prasarana selalu berada pada posisi pertama sebagai aspek yang paling banyak dibicarakan.



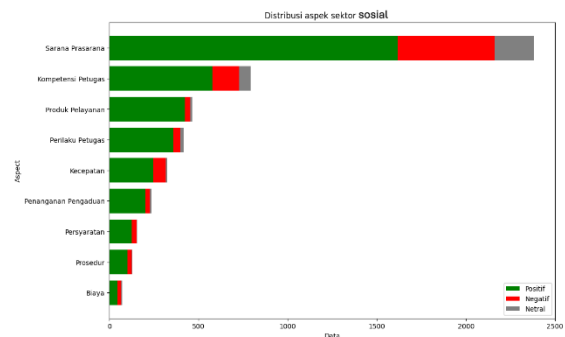
Gambar 6. Distribusi aspek



Gambar 7. Distribusi aspek sektor transportasi



Gambar 8. Distribusi aspek sektor kesehatan



Gambar 9. Distribusi aspek sektor sosial

Dari hasil proses pelabelan dan ekstraksi aspek ini, kemudian dijadikan sebagai data latih dan *fine-tuning* pada model IndoBERT. Hasil evaluasi pada Tabel I membandingkan nilai *eval\_loss* di tiga variasi pembagian data pelatihan dan validasi: 80:20, 70:30, dan 60:40, masing-masing dengan 3 *epoch*. Hasilnya

menunjukkan bahwa rasio 80:20 menghasilkan *eval\_loss* terendah sebesar 0,1804, yang menandakan performa model paling optimal dengan pembagian ini. Pada rasio 70:30, *eval\_loss* sedikit meningkat menjadi 0,1897, tetapi tetap relatif rendah. Sementara itu, pada rasio 60:40, *eval\_loss* meningkat secara signifikan hingga mencapai 0,9447, yang mengindikasikan bahwa pengurangan data pelatihan berdampak negatif terhadap kemampuan model untuk melakukan generalisasi dengan baik terhadap data validasi.

Pada rasio 80:20, model mencapai akurasi tertinggi sebesar 95,54%, dengan presisi sebesar 95,53%, yang mengukur ketepatan prediksi positif. *Recall*, yang mengukur sejauh mana model mendeteksi seluruh data positif, juga tinggi pada rasio ini (95,54%). Pada rasio 70:30, meskipun terdapat sedikit penurunan, hasilnya tetap cukup konsisten, dengan akurasi sebesar 95,36% dan presisi serta *recall* masing-masing sebesar 95,35% dan 95,36%. Namun, pada rasio 60:40, terdapat penurunan yang lebih signifikan, dengan akurasi turun menjadi 94,66% serta presisi dan *recall* turun menjadi 94,47% dan 94,55%. Hal ini menunjukkan adanya penurunan kinerja model seiring dengan berkurangnya data pelatihan.

Tabel 1. Evaluasi komparasi data latih

Parameter	Komparasi data latih		
	80:20	70:30	60:40
<i>eval_loss</i>	0.1804	0.1897	0.9447
<i>eval_accuracy</i>	95,54%	95,36%	94,66%
<i>eval_precision</i>	95,53%	95,35%	94,47%
<i>eval_recall</i>	95,54%	95,36%	94,55%
<i>eval_f1</i>	95,35%	95,36%	94,47%

Hasil dari skenario pertama kemudian diterapkan dalam skenario kedua dengan menggunakan beberapa sampel data dari berbagai sektor. Seperti yang ditunjukkan pada Tabel II, metrik evaluasi diperoleh dari tiga sektor sosial, transportasi dan kesehatan. Nilai *eval\_loss* menunjukkan seberapa baik kinerja model, dengan angka yang lebih rendah menunjukkan hasil yang lebih baik. Sektor sosial memiliki nilai *loss* terendah sebesar 0,2960, diikuti oleh sektor transportasi dengan nilai 0,2766, dan sektor kesehatan memiliki nilai *loss* tertinggi sebesar 0,3663. Hal ini menunjukkan bahwa model bekerja lebih baik di sektor sosial dan transportasi dibandingkan dengan sektor kesehatan, di mana model mengalami lebih banyak kesulitan.

Tabel 2. Evaluasi komparasi sampel data

Parameter	Sampel data sektor		
	Sosial (12146 data)	Transp- ortasi (10769 data)	Kesehatan (8871 data)
<i>eval_loss</i>	0.2960	0.2766	0.3663
<i>eval_accuracy</i>	93,08%	92,47%	90,87%
<i>eval_precision</i>	92,97%	92,44%	90,66%
<i>eval_recall</i>	93,08%	92,47%	90,87%
<i>eval_f1</i>	93,01%	92,45%	90,72%

#### 4. KESIMPULAN

Analisis sentimen berbasis aspek menggunakan model IndoBERT dapat dianggap menghasilkan performa yang baik, terutama pada data survei kepuasan publik dalam aplikasi E-SKM milik Pemerintah Provinsi Jawa Tengah. Berdasarkan hasil evaluasi model IndoBERT untuk tugas analisis sentimen, proporsi data pelatihan dan uji, maupun jumlah data di setiap sektor, memiliki dampak signifikan terhadap kinerja model. Model ini menunjukkan performa terbaik ketika menggunakan proporsi data pelatihan dan uji 80:20, dengan nilai *eval\_loss* terendah (0,1804) dan akurasi tertinggi (95,54%). Ketika proporsi data pelatihan menurun, seperti pada pembagian 60:40, performa model sedikit menurun, dengan *eval\_accuracy* yang lebih rendah (94,66%) dan *eval\_loss* yang lebih tinggi (0,9447). Selain itu, evaluasi pada berbagai sektor data (sosial, transportasi, dan kesehatan) menunjukkan bahwa jumlah data juga memengaruhi kinerja model. Sektor sosial, yang memiliki dataset terbesar (12.146 data), menghasilkan performa terbaik dengan *eval\_accuracy* tertinggi sebesar 93,08% dan skor *eval\_f1* sebesar 93,01%. Sebaliknya, sektor kesehatan, dengan dataset terkecil (8.871 data), menunjukkan performa yang lebih rendah, dengan *eval\_accuracy* sebesar 90,87%. Oleh karena itu, selain proporsi data pelatihan yang lebih besar, memiliki jumlah data yang cukup juga penting untuk meningkatkan akurasi dan performa keseluruhan model.

Dari hasil ini dapat ditarik kesimpulan bahwa model IndoBERT memiliki kelebihan pada tugas dengan jumlah data yang besar dilihat dari parameter evaluasi pada masing-masing sampel data di Tabel 2. Dari tabel ini juga dapat dilihat kelemahan model IndoBERT yaitu terjadi penurunan performa pada masing-masing parameter evaluasi seiring jumlah data yang semakin kecil. Namun, hal ini tidak menutup kemungkinan adanya hasil yang berbeda jika diterapkan pada data yang berbeda. Sebagai bahan referensi untuk penelitian selanjutnya, kesimpulan ini tidak menutup kemungkinan memperoleh hasil yang berbeda jika model ini digunakan dalam tugas lain yang berbeda seperti mendeteksi spam, hoax dan tugas pada pengolahan bahasa alami yang lainnya. Selain itu, untuk mengetahui hasil performa yang lebih komperhensif pada model IndoBERT dapat dilakukan pada penelitian selanjutnya dengan menggunakan kamus leksikon yang berbeda atau variasi pada *preprocessing* data, hal ini dapat mengetahui selain jumlah data, apakah penggunaan kamus leksikon dan *preprocessing* juga memiliki pengaruh pada model ini.

Ucapan terima kasih disampaikan kepada semua pihak yang telah memberikan dukungan, masukan, serta data yang diperlukan selama proses penelitian ini berlangsung. Penulis berharap penelitian ini

memiliki kontribusi keilmuan pada analisis sentimen berbasis aspek dan penggunaan model IndoBERT.

#### DAFTAR PUSTAKA

- ALQARYOUTI, O., SIYAM, N., dan SHAALAN, K. 2024. Aspect-based sentiment analysis for government smart applications customers' reviews. *Applied Computing and Informatics*, 20(1/2), 142-161.
- ASSIRI, A., GUMAEI, A., MEHMOOD, F., ABBAS, T., dan ULLAH, S. 2024. DeBERTa-GRU: Sentiment Analysis for Large Language Model. *Computers, Materials & Continua*, 79(3), 4219-4236.
- BRAUWERS, G., dan FRASINCAR, F. 2021. A survey on aspect-based sentiment classification. *ACM Computing Surveys*, 54(3), 1-37.
- CHAUDHARY, A., PANDEY, D. K., dan AGRAWAL, A. 2022. Aspect-based sentiment analysis: A survey of the state-of-the-art. *Expert Systems with Applications*, 204, 117547.
- CITRA, B. S. A., SETIONO, B., PANGARIBUAN, C. H., dan AMBARWATI, M. F. L. 2021. The Influence of Service Quality on Public Satisfaction and Public Trust: A Study on Jakarta Public Health Services during COVID-19 Pandemic. *Journal of Business, Management, and Social Studies*, 1(1), 48-57.
- DEVLIN, J., 2018. BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint*, arXiv:1810.04805.
- DIEKSON, Z. A., PRAKOSO, M. R. B., PUTRA, M. S. Q., SYAPUTRA, M. S. A. F., ACHMAD, S., dan SUTOYO, R. 2023. Sentiment analysis for customer review: Case study of Traveloka. *Procedia Computer Science*, 216, 682-690.
- ELIFNEH, Y. W., BRAHMA, D., JAGADISH, G., dan GIRMA, Y., 2020. Customers' satisfaction in ATM service: empirical evidence from the leading bank in Ethiopia. *International Journal of Engineering and Management Research*, 10(1), pp. 126-134. GUEROLA NAVARRO, V., OLTRA-BADENES, R., GIL GOMEZ, H., and GIL-GOMEZ, J. A., 2021. Research model for measuring the impact of customer relationship management (CRM) on performance indicators. *Economic Research-Ekonomika Istraživanja*, 34(1), pp. 2669-2691.
- GU, T., HE, Z., LI, Z., dan WAN, Y. 2025. Information-assisted and sentiment relation-driven for aspect-based sentiment analysis. *Expert Systems with Applications*, 278, 127308.
- HARIRI, R. H., FREDERICKS, E. M. dan BOWERS, K. M., 2019. Uncertainty in big data analytics: survey, opportunities, and challenges. *Journal of Big Data*, 6(1).
- HIDAYAT, A., SAPUTRA, R., dan RAHMAN, F. 2023. Analisis Pengaruh Penerapan Stopword Removal pada Performa Klasifikasi Sentimen Tweet Bahasa Indonesia. *Jurnal Teknologi dan Sistem Komputer*, 11(3), 235-242.
- IBNU DAQIQIL, H., SAPUTRA, H., SYAMSUDHUHA, KURNIAWAN, R. dan ANDRIYANI, Y., 2024. Sentiment analysis of student evaluation feedback using transformer-based language models. *Indonesian Journal of Electrical Engineering and Computer Science*, 36(2), pp. 1127-1139 (pp. 168-177).
- JADHAV, A., dan SONAR, P. 2021. Aspect based sentiment analysis datasets: A comparative study. *Procedia Computer Science*, 185, 578-587.
- KOTO, F., WINATA, G. I., SIDI, R., dan MARDIYANTO, R., 2020. IndoLEM and IndoBERT: A benchmark dataset and pre-trained language model for Indonesian NLP. *arXiv preprint*, arXiv:2011.00677.
- MA, Y., PENG, H., dan CAMBRIA, E. 2022. Targeted aspect-based sentiment analysis via embedding commonsense knowledge graph. *IEEE Transactions on Affective Computing*.
- NABIILAH, G. Z., PRASETYO, S. Y., dan IZDIHAR, Z. N. 2023. BERT base model for toxic comment analysis on Indonesian social media. *Procedia Computer Science*, 216, 714-721.
- PINTO LUQUE, E. B., 2023. An integrated NLP approach to sentiment analysis in satisfaction surveys. *arXiv preprint*, arXiv:2307.11771.
- ROMERO SUBIA, J. F., JIMBER-DEL RIO, J. A., OCHOA-RICO, M. S. dan VERGARA-ROMERO, A., 2022. Analysis of citizen satisfaction in municipal services. *Economies*, 10(9), p. 225.
- SELVAKUMAR, B., dan LAKSHMANAN, B. 2022. Sentimental analysis on user's reviews using BERT. *Materials Today: Proceedings*, 62, 4931-4935.
- ZHANG, W., LI, X., DENG, Y., BING, L. dan LAM, W., 2022. A survey on aspect-based sentiment analysis: Tasks, methods, and challenges. *IEEE Transactions on Knowledge and Data Engineering*, 35(11), pp. 11019-11038.

ZIEMSKI, M., CHAURASIA, S., dan SCHERER, R.  
2020. Aspect extraction in sentiment  
analysis: Comparative analysis and survey.

*Artificial Intelligence Review*, 53, 2357-  
2390.